MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 30, 1998	Final 3 March 95 - 2 December 97

4. TITLE AND SUBTITLE

Novel Approaches to the Characterization of Specific Protein-Protein Interactions. Important in Gene Expression

5. FUNDING NUMBERS

DAAH04-95-1-0138

6. AUTHOR(S)

Ronald L. Somerville

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Purdue Research Foundation
Division of Sponsored Programs
Hovde Hall
West Lafayette, IN 47907

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 33688.1-LS

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Through the application of a range of techniques in the areas of biochemistry, molecular biology, and microbial physiology, we have purified and characterized the TyrR protein of *Haemophilus influenzae*. This dimeric protein interacts with specific operator targets associated with promoters that drive the production of proteins essential for aromatic amino acid biosynthesis or transport. Like its *E. coli* counterpart, the *H. influenzae* protein is organized into discrete domains that bind either DNA or small molecules that function as coregulators. In separate studies, we showed that the tyrosine phenol lyase-promoter of *Citrobacter freundii* is regulated not only by the TyrR protein but also by two global transcription factors, namely Integration Host Factor and cyclic AMP receptor protein.

14. SUBJECT TERMS

Helix-turn-helix motif; transcriptional regulation, TyrR protein; *Escherichia coli*; tyrosine degradation

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std 239-18
298-102

Novel Approaches to the Characterization of Specific Protein-Protein Interactions Important in Gene Expression

Final Technical Report

Ronald L. Somerville, Ph.D.

October 30, 1998

U.S. Army Research Office

Grant No. DAAH 04-95-1-0138

Purdue University
West Lafayette, IN

A. STATEMENT OF PROBLEM

(i) The TyrR regulon. In *E. coli*, there are eight different transcriptional units whose expression is controlled either positively or negatively by the TyrR protein (reviewed by Pittard, 1996). Recently, we have shown that the *tpl* gene of *Citrobacter freundii* is positively controlled by TyrR (see below). Each transcriptional unit encodes one or more proteins involved in either the biosynthesis, transport, or degradation of the three aromatic amino acids, tyrosine, phenylalanine and tryptophan. These amino acids can modulate one or more of the activities of the TyrR protein.

The TyrR protein of *E. coli* is a homodimer containing 513 amino acid residues per monomer (Cornish *et al.* 1986; Yang *et al.*, 1993). In limited proteolysis studies (Cui and Somerville, 1993a) it was shown that TyrR has two identifiable domains, consisting of residues 1-190 and 191-467. Residues 467-513 do not form a discrete domain, and contain the helix-turn-helix DNA-binding motif.

The second domain of TyrR (residues 191-467) has a high degree of homology to the central domains of a number of transcriptional activators belonging to the NtrC-NifA superfamily (Stock *et al.*, 1989). In general, such proteins activate transcription from σ^{54} promoters, exert their action from a distance, and have a modular structure (Kustu *et al.*, 1991; Morett and Segovia, 1993). In contrast, TyrR affects transcription only from σ^{70} promoters. Many of the transcription factors in the NtrC superfamily are members of two-component systems that acquire regulatory potency upon phosphorylation by a related histidine protein kinase (Stock *et al.*, 1995). However, TyrR controls transcription without the collaboration of any other protein component, in response to the binding of an effector amino acid such as phenylalanine or tyrosine. In this respect TyrR resembles a rapidly growing mechanistic subgroup of the NtrC superfamily of proteins that directly senses and responds to small effector molecules.

(ii) The TyrR protein of *Haemophilus influenzae*: a "mini-member" of the NtrC superfamily. A major stimulus to TyrR research was the elucidation by Fleischmann and co-workers (1995) of the complete genomic sequence of *H. influenzae* Rd. Within this organism there was only one predicted ORF (HI 0410) that matched the consensus sequence for the central domain of the NtrC superfamily of transcription factors. In their annotation of the *H. influenzae* sequence, Fleischmann *et al.*, called HI 0410 as the structural gene for TyrR, largely because there was an excellent match at the C-termini, where the amino acid residues responsible for operator recognition lie. Surprisingly, the *H. influenzae* protein (hereinafter designated $TyrR_{HI}$) contained only 318 amino acid residues. Essentially all of the difference in size could be accounted for by assuming that $TyrR_{HI}$ lacked a homologue of the N-terminal activation domain of the *E. coli* protein that had been described in earlier work (Cui and Somerville, 1993a, b). The validity of these predictions was evaluated in biochemical and genetic studies, described below.

B. SUMMARY OF MOST IMPORTANT RESULTS

1. Characterization of the TyrR Protein of *H. influenzae*. The gene that was inferred to encode the TyrR protein of *Haemophilus influenza*e Rd was synthesized by polymerase chain reaction and inserted into a T7-based expression vector. Methods were developed to overexpress

the TyrR protein of *H. influenzae* in *Escherichia coli* and to purify the protein on a large scale. Both *in vitro* and *in vivo* functional comparisons of the *H. influenzae* and *E. coli* TyrR proteins were carried out. The TyrR protein of *H. influenzae* was able to bind *in vitro* to an operator target upstream of the *aroF-tyrA* gene of *E. coli*. In the presence of [γ-S]-ATP, the DNA-binding ability of the *H. influenzae* TyrR protein was drastically reduced. Despite the much shorter peptide chain length (318 amino acid residues *vs.* 513), the TyrR protein of *H. influenzae* was as active in repressing the *aroF* promoter as the TyrR protein of *E. coli*. Repression by both proteins was enhanced in the presence of tyrosine; however, the transcriptional activation function associated with the TyrR protein of *E. coli* could not be detected when the *H. influenzae* TyrR protein was expressed in *E. coli*. By computer analysis, at least five operator targets for TyrR were identified within the genomic DNA of *H. influenzae*. These observations show that the assignment of function to the *tyrR* gene of *H. influenzae* was correctly made.

Highly purified preparations of the TyrR protein of *Haemophilus influenzae Rd* undergo specific and limited proteolytic cleavage during storage at 4°C to generate two fragments of 28 kDa and 8 kDa. Under nondenaturing conditions, the two fragments remain tightly associated. Nicked TyrR is identical to full-length TyrR in its operator binding characteristics. The 8 kDa fragment, containing amino acid residues 258-318, was separated from the 28 kDa fragment (residues 1-257) by gel filtration chromatography in the presence of 4 M urea. Upon renaturation, this fragment bound to operator with an affinity similar to that of full length TyrR, but was unresponsive to ligands that normally modulate operator binding (γ-S-ATP and L-tyrosine). It was not possible to renature the urea-treated 28 kDa fragment. Highly purified soluble preparations of truncated TyrR, containing residues 1-257, were obtained following the overexpression of a shortened form of the *tyrR* gene via a specific plasmid construct. By several criteria, this species had native secondary and tertiary structure. The 28 kDa fragment was unable to bind to operator but could reconstitute nicked TyrR when added to the renatured 8 kDa fragment, as shown by physical properties and responsiveness to cofactors in operator binding. When either the 28 kDa or 8 kDa species was expressed *in vivo*, there was no detectable operator binding, as evaluated using a *lacZ* reporter system driven by the repressible *aroF* promoter. When the two fragments were co-expressed in a common cytoplasm, an operator-binding species was formed, as demonstrated through partial restoration of repression capability.

2. Studies on the TyrR Protein of *Escherichia coli*. The TyrR protein is the chief transcriptional regulator of a group of genes in *Escherichia coli* whose protein products either catalyze reactions in aromatic amino acid biosynthesis or mediate the transport of aromatic amino acids. The affinity of TyrR for an operator target within the *aroF* promoter was reduced more than 40-fold by Zn^{2+} (100 μM). The addition of L-tyrosine and either ATP or ADP restored operator binding activity to Zn^{2+}-liganded TyrR. The affinity of Zn^{2+}-liganded TyrR for operator increased by about 200-fold in the presence of γ-[S]-ATP and L-tyrosine. In the absence of Zn^{2+}, these ligands mediated only a 5-fold increase in operator affinity. All of the known catalytic activities of TyrR were profoundly altered by Zn^{2+}. At 100 μM Zn^{2+}, the ATPase activity of TyrR was strongly inhibited. Under the same conditions, the ability of TyrR to self-phosphorylate and to hydrolyze p-nitrophenyl phosphate was stimulated. A 15% decrease in fluorescence intensity was observed upon interaction of Zn^{2+} (400 μM) with metal-free TyrR protein.

3. Cloning and characterization of the *tyrR* Genes of *Salmonella typhimurium* and *Citrobacter braakii*. The TyrR protein of *Escherichia coli* is a transcription factor with substantial sequence homology to NtrC and numerous other proteins that modulate σ^{54} promoters. However, TyrR is specific for σ^{70} promoters. To expand the information base of TyrR sequences, the *tyrR* genes of *Citrobacter braakii* and *Salmonella typhimurium* were synthesized by PCR, cloned and structurally analyzed. The DNA sequence of the *tyrR* gene of *C. braakii* predicts a protein of 514 amino acids, while a protein of 513 amino acids is predicted from the *tyrR* gene of *S. typhimurium*. The newly characterized TyrR proteins are 80% identical to the TyrR protein of *E. coli*. Those amino acid residues of *E. coli* TyrR that are known to be functionally important were all conserved in the TyrR proteins of *C. braakii* and *S. typhimurium*. When the central domain of each known TyrR protein is compared to other members of the NtrC superfamily of transcription factors, a unique gap of eight amino acid residues is found. This gap is situated at residues 282/283 of full-length TyrR. It is predicted that this gap is required for interaction between the σ^{70} RNA polymerase and TyrR.

4. Studies of Transcriptional Control at the *tpl* Promoter. The ability of microorganisms to degrade L-tyrosine to phenol, pyruvate and ammonia is catalyzed by the inducible enzyme L-tyrosine phenol-lyase (EC 4.1.99.2). To investigate possible mechanisms for how the synthesis of this enzyme is regulated, a variety of biochemical and genetic procedures were used to analyze transcription from the *tpl* promoter of *Citrobacter freundii* ATCC 29063 (*C. braakii*). By computer analysis of the region upstream of the *tpl* structural gene, two segments of DNA bearing strong homology to the known operator targets of the TyrR protein of *Escherichia coli* were detected. A DNA fragment of 509 bp carrying these operator targets plus the presumptive *tpl* promoter was synthesized by PCR and used to construct a single-copy *tpl-lacZ* reporter system. The formation of β-galactosidase in strains carrying this reporter system, which was measured in *E. coli* strains of various genotypes, was strongly dependent on the presence of a functional TyrR protein. Several mutationally altered forms of TyrR were deficient in their abilities to activate the *tpl* promoter. The pattern of loss of activation function was exactly parallel to the effects of the same *tyrR* mutations on the *mtr* promoter, which is known to be activated by the TyrR protein. When cells carrying the *tpl-lacZ* reporter system were grown on glycerol, the levels of β-galactosidase were 10-20 fold higher than those observed in glucose-grown cells. The effect was the same whether or not TyrR-mediated stimulation of the *tpl* promoter was in effect. By deleting the *cya* gene, it was shown that the glycerol effect was attributable to stimulation of the *tpl* promoter by the cyclic AMP (cAMP) receptor protein system. A presumptive binding site for this transcription factor was detected just upstream of the -35 recognition hexamer of the *tpl* promoter. The transcriptional startpoint of the *tpl* promoter was determined by chemical procedures. The precise locations of the TyrR binding sites, which were established by DNase I footprinting, agreed with the computer-predicted positions of these regulatory sites. The two TyrR operators, which were centered at coordinates -272.5 and -158.5 with respect to the transcriptional start point, were independently disabled by site-directed mutagenesis. When the upstream operator was altered, activation was completely abolished. When the downstream operator was altered, there was a fourfold reduction in reporter enzyme levels. The *tpl* system presents a number of intriguing features not previously encountered in TyrR-activated promoters. First among these is the question of how the TyrR protein, bound to widely separated operators, activates the *tpl* promoter which is also widely separated from the operators.

The *tpl* gene of *Citrobacter freundii* encodes an enzyme that catalyzes the conversion of L-tyrosine to phenol, pyruvate, and ammonia. This gene is known to be positively regulated by TyrR. The amplitude of regulation attributable to this transcription factor is at least 20-fold. Three TyrR binding sites, designated Boxes A, B, and C, centered at coordinates -272.5, -158.5, and -49.5, respectively, were identified in the upstream region of the *tpl* promoter. The results of mutational experiments suggest that TyrR binds in cooperative fashion to these sites. The nonavailability of any TyrR site impairs transcription. Full TyrR-mediated activation of *tpl* required integration host factor (IHF) and the cAMP receptor protein (CRP). By DNase I footprinting, it was shown that the IHF binding site is centered at coordinate -85 and that there are CRP binding sites centered at coordinates -220 and -250. Mutational alteration of the IHF binding site reduced the efficiency of the *tpl* promoter by at least eighhtfold. The proposed roles of CRP and IHF are to introduce bends into *tpl* promoter DNA between boxes A and B or B and C. Multimeric TyrR dimers were demonstrated by a chemical cross-linking method. The formation of hexameric TyrR increased when *tpl* DNA was present. The participation of both IHF and CRP in the activation of the *tpl* promoter suggests that molecular mechanisms quite different from those that affect other TyrR-activated promoters apply to this system.

C. PUBLICATIONS AND TECHNICAL REPORTS

Chen, Y.-W., Dekker, E.E. and Somerville, R.L. (1995) Functional analysis of *E. coli* threonine dehydrogenase by means of mutant isolation and characterization. *Biochim. Biophys. Acta* **1253**, 208-214.

Zhu, Q., Zhao, S. and Somerville, R.L. (1997) Expression, purification, and functional analysis of the TyrR protein of *Haemophilus influenzae*. *Protein Expression and Purification* **10**, 237-246

Smith, H.Q. and Somerville, R.L. (1997) The *tpl* promoter of *Citrobacter freundii* is activated by the TyrR protein. *J. Bacteriol.* **179**, 5914-5921.

Bai, Q. and Somerville, R.L. (1998) Integration host factor and cyclic AMP receptor protein are required for TyrR-mediated activation of *tpl* in *Citrobacter freundii*. *J. Bacteriol.* **180**, in press.

D. PARTICIPATING SCIENTIFIC PERSONNEL

1. Ihor Skrypka (Graduate Research Assistant, 3 March 95 - 31 October 96). Awarded Ph.D. December 1996.

2. Qing Bai (Graduate Research Assistant, 3 March 95 - 2 December 97). Awarded Ph.D. August 1998.

3. Shimin Zhao (Graduate Research Assistant, 3 March 95 - 2 December 97).

4. Liuying Ni (Professional Research Associate, 3 March 95 - 2 December 97).

5. Hong Qiu Smith (Professional Research Associate, 3 March 95 - 2 December 97). Awarded M.S. May 1998.

6. Karen Qiu Zhu (Postdoctoral Research Fellow, 19 November 95 - 19 November 96).

BIBLIOGRAPHY

Cornish, E.C., Argyropoulos, V.P., Pittard, J. and Davidson, B.E. (1993) Structure of the *Escherichia coli* K12 regulatory gene *tyrR*. *J. Biol. Chem.* **261**, 403-410.

Cui, J. and Somerville, R.L. (1993a) The TyrR protein of *Escherichia coli*. Analysis by limited proteolysis of domain structure and ligand-mediated conformational changes. *J. Biol. Chem.* **268**, 5040-5047.

Cui, J. and Somerville, R.L. (1993b) A mutational analysis of the structural basis for transcriptional activation and monomer-monomer interaction in the *tyrR* system of *Escherichia coli* K-12. *J. Bacteriol.* **175**, 1777-1784.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C. (1995) Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.

Kustu, S., North, A.K. and Weiss, D.S. (1991) Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends Biochem. Sci.* **16**, 397-402.

Morett, E. and Segovia, L. (1993) The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* **175**, 6067-6074.

Pittard, A.J. (1996) Biosynthesis of the aromatic amino acids. In *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology, ed. by F.C. Neidhardt. ASM Press, Washington. pp 458-484.

Stock, J.B., Ninfa, A.J. and Stock, A.M. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **53**, 450-490.

Stock, J.B., Surette, M.G., Levit, M. and Park, P. (1995) Two-component signal transduction systems: Structure-function relationships and mechanisms of catalysis. In *Two-Component Signal Transduction*, ed. by J.A. Hoch and T.J. Silhavy. ASM Press, Washington. pp. 25-51.

Yang, J., Camakaris, H. and Pittard, A.J. (1993b) Mutations in the *tyrR* gene of *Escherichia coli* which affect TyrR-mediated activation but not TyrR-mediated repression. *J. Bacteriol.* **175**, 6372-6375.